Assessment of Nutrient Loading and Eutrophication in Barnegat Bay-Little Egg Harbor, NJ in Support of Nutrient Management Planning

Peer Review Scope of Work

Key Topics for Consideration:

Data sufficiency to draw conclusions/develop index – There are significant limitations with respect to the data available for this project – examples include a lack of data for some measures across the years considered, limited frequency of data collection (typically quarterly monitoring so that if growing season is selected as stated, there may be only two sampling events), and limited data available for some locations. Given these limitations, can the conclusions drawn and/or the index developed be used for management purposes, developing strategies to target specific levels of nutrients that would be expected to result in support of a healthy ecosystem for this estuary? Are conclusions supported by the data? Specific examples of data limitations are listed below:

1. Data quality concerns: Were censored values, non-detects, zero values, skewness, outliers handled correctly/adequately?
2. Are the written conclusions in line with data presented? Should there be any concerns regarding poor statistical correlations?
3. The study states that, "The BB-LEH database was analyzed for each segment of the bay, because these segments have been determined to be heterogeneous habitats." If this statement is true, is the determination of one threshold calculation for the entire bay for each indicator appropriate in determining the indicator score for each segment, or should the threshold calculations for indicators be defined separately for each segment in order to determine the indicator score for each segment of the Bay?
4. Do current USGS studies sufficiently capture (identify and estimate) all substantive N and P loads to the bay? If not, please identify additional sources that should be considered.
5. Do the included condition variables include all important parameters of interest regarding the bay's condition? Is it important or useful to have any estimates of microbial loop or secondary production (e.g., if only for the bay's herbivores [clams])?
6. Given the methodology used to derive a unit-less score for the index, the index assessment for any given year is opportunistic (limited by the data available for a given year) and not deterministic (informed by data from the full suite of prospective relevant factors). As a result, the importance of setting thresholds against which observations are compared to determine the assessment cannot be overstated. As the value for each threshold is one of the most important elements in determining the outcome of applying the index, it is essential that the threshold values be solidly based in science. In other estuary studies, the causal thresholds (for nutrients N and P) were selected based on modeling the relationship between the causal factor and the response variables in the particular waterbody, which is appropriate because the fate and transport of nutrients will vary given the physical/chemical/biological dynamics unique to that water body. Here, the causal thresholds were selected before that modeled relationship has been determined. Does this limit the study's use for management purposes, developing strategies to target specific levels of N and P that would be expected to result in support of a healthy ecosystem for this estuary?

Selection of thresholds to define condition – The basis for selecting the thresholds is given as literature, data analysis, best professional judgment (BPJ) or a combination of these factors. Is this a supportable basis for selecting thresholds that would be used to make condition assessments and inform management options designed to effectuate improvement in condition? Specific concerns and questions:

1. Is there sufficient information within the study report to show that there is enough Barnegat Bay data to determine each of the threshold indicator values? Has the report addressed which indicators relied more heavily on BPJ or literature and should be revisited when more Barnegat Bay specific data for that indicator are available? Is the report detailed and transparent enough such that the reader can reproduce all steps taken to get to the conclusions provided?
2. Are you aware of any other significant data/studies that are relevant and should be included or referenced in this study and should have been used to help determine the threshold indicator values? Please explain fully.

Derivation of the index – The derivation of the index relies on a PCA analysis and a series of manipulations involving raw data values and weighted scores. Is the derivation of the index in the manner indicated supportable?
Specific concerns:

3. Determination of index values blends raw scores (comparison of average of raw data to a selected threshold) and weighted scores (square of eigenvector value, considering the factors for which there was data in a given year). Weighted scores simply represent a measure of the variability of the factor, if it is present within a given year. If there are no data, the factor is given no weight. What is the purpose of blending the weighted score with the raw score, and is this a valid approach?
4. The approach taken in using PCA is not standard and no documentation is presented to justify it. Typically, to develop an index using PCA, the scores of the first few principal components would be examined. If the first eigenvalue (score variance) comprises a large amount of the total variability, then the first principal component might be taken as the index. If weighting the index is desired then the first eigenvalue would be used as a weight. In this report, there do not seem to be any attempts to assess the adequacy of using only the first principal component. Is this approach valid? If not, what argument, further analysis, and documentation would justify this approach?
5. The approach taken in this report is to use the squared component of the eigenvector as a multiplicative weight for that component of the index. The justification is that this weight would be the variance of the component. Is this claim correct? If the variables had been standardized to a variance of 1, then there would be some basis for this, although correlations between variables would also have to be considered. The SAS code in the appendices shows that no variance standardization was done during the PCA analysis and it did not appear to have been done before that. Should the use of multiplicative weighting not be justified, as well as this particular weighting method? Do these concerns affect the validity of the index's derivation, and what can be done to address them?
6. The sole justification for combining the weighted and raw indices is that it integrates the multiple indicators and their variability. The advantage of this approach is not obvious and requires some justification and documentation. Would combining the two indices serve to blur any useful measure, or instead improve it? Do these concerns affect the validity of the index measures?

Use of Index

1. Objective 5 of this study is "To generate an Index of Eutrophication as a tool to evaluate future conditions using water quality and biotic indicators to assess eutrophication, eutrophic impacts, and overall ecosystem health of the BB-LEH Estuary..." Does the study report provide enough information on how one can use the Index of Eutrophication to evaluate future conditions using newly acquired water quality data? Is the report detailed and transparent enough such that the reader can reproduce all sets taken to get to the conclusions provided?
2. In your opinion, what are the weakest and the strongest aspects of the Eutrophication Index and the Threshold determinations? Please make suggestions on how the weakest parts can be strengthened.
3. Are there any elements missing from the Eutrophication Index which you think need to be included or which would strengthen the tool? Please explain fully.
4. The Estuaries and Coasts article Mind the Data Gap: Identifying and Assessing Drivers of Changing Eutrophication Condition (Fertig, et al.) identifies grouping the variables into three major categories, one of which is seagrass, to develop an index of eutrophication; however, there are no seagrass data available for the first 15 of the 25 years of data used to develop the index. Thus, can we be confident in using and applying this index?
5. Does the approach used here "validate" the developed eutrophication index?

Overall Adequacy of Report

1. Is the organization of the document appropriate and does it present the material in a clear and concise manner? Please explain fully.
2. Are the stated objectives adequately met? Please explain fully.
3. Do the results from the study support the authors' conclusions and recommendations?